

FIGURE 1A

Frame3 L C S L L F G L S L S I T M F E M M P
 Frame2 V M * S F V W S L P * H N D V * D D A I
 Frame1 C Y V V F C L V S P L A * R C L R * C H
 DNA TGTTATGTAGTCCTTGTGGTCTCCCTTAGCATAACGATGTTGAGATGATGCCAT 60
 K hcv16029210 rpt

.....
 Frame3 F I H F C C * A A A E Y C W N P S L F I
 Frame2 H S F L L L S S C R V L L E S Q F I H W
 Frame1 S F I F V A E Q L P S I V G I P V Y S L
 DNA TCATTCACTTTGTTGCTGAGCAGCTGCCAGTATTGTTGGAATCCCAGTTATTCAATTG 120
 Y hcv26544776 SNPG98 AL662833

.....
 Frame3 G F C V S S * * T C G F L Q L G F V I N
 Frame2 F L C L Q L I D M W I P P V R V C Y * *
 Frame1 V S V S P V D R H V D S S S * G L L L M
 DNA GTTCTGTGTCCTCAGTTGATAGACATGTGGATTCCCTCAGTTAGGGTTATTAAATG 180
 Y hcv26544777 rpt
 | SNPG168 AL662833

.....
 Frame3 E A T I N N C L Q V W T Y I F I S F G *
 Frame2 S H Y K * L L T S V D L H F Y F F W I N
 Frame1 K P L * I T A Y K C G L T F L F L L D K
 DNA AAGCCACTATAAAATAACTGCTTACAAGTGTGGACTTACATTTTATTCTTTGGATAAA 240
 R hcv27464285 rpt
 | SNPG205 AL67188

.....
 Frame3 I R I C G I A G P C G N R W V T V * E T
 Frame2 T Y L W N C W A M W * * M G N C I R N C
 Frame1 Y V F V E L L G H V V I D G * L Y K K L
 DNA TACGTATTTGTTGGAATTGCTGGCCATGTGGTAATAGATGGTAACTGTATAAGAAACTG 300
 hcv26544778 Y rpt
 | SNPG298
 AL662833

.....
 Frame3 A I P L Y K L A A T F F A F L P A I S D
 Frame2 H T T L Q I G C H I F C I P T S N I R H
 Frame1 P Y H F T N W L P H F L H S Y Q Q Y Q T
 DNA CCATACCACTTACAATTGGCTGCCACATTTCCTACCAGCAATATCAGACA 360
 hcv27464284 K rpt
 | SNPG354

AL67188

Y hcv27465835 rpt
 | SNPG333 AL67188

.....
 Frame3 I P I F S I F L P V L R L I I C L F N F
 Frame2 S Y F F H I L A S V K T Y H M S F * L Y
 Frame1 F L F F P Y S C Q C * D L S Y V F L T L
 DNA TTCCCTATTTTCCATATTCTTGCCAGTGTAAAGACTTATCATATGTCTTTAACCTTA 420

.....
 Frame3 I C S R * C V M V S H C G F N L H F F D
 Frame2 L L * V M C D G F S L W F * L A L L * *
 Frame1 S A L G D V * W F L I V V L T C T S L M
 DNA TCTGCTCTAGGTGATGTGATGGTTCTCATTGTGGTTAACCTGCACCTTTGATG 480
 Y hcv27466112 rpt
 | SNPG421 AL67188

| w SNPG476
 AL67188

2/12

FIGURE 1B

Frame3 D * Y C L L S F H V H L S D L L L H I F Y
 Frame2 L V L F A I F S C S S K R L I T Y I L *
 Frame1 T S I V C Y L F M F I * A T Y Y I Y F M
 DNA ACTAGTATTGTTGCTATCTTTCATGTTCATCTAAGCGACTTATTACATATATTTATG 540
 hcv27465418 Y | SNPG518 AL67188 | SNPG539

 Frame3 E L F C K F N D * F Q R L F Q N S L V F
 Frame2 T I L Q I Q * L I P E T F S E F P S V F
 Frame1 N Y F A N S M I N S R D F F R I P * C F
 DNA AACTATTTGCAAATTCAATGATTAATTCCAGAGACTTTCTAGTGTTT 600

 Frame3 S T Y T M K L V T K K D F H F F L S Y P
 Frame2 Y I Y N E V G D K E R L S F L P F L S I
 Frame1 L H I Q * S W * Q R K T F I S S F L I H
 DNA CTACATATACAATGAAGTTGGTGACAAAGAAAGACTTTCTTCCTTCTTATCCAT 660
 | Y SNPG618 hcv26544780AL662833 AL67188

 Frame3 L I F F L L K L L F G R D E V S L I R
 Frame2 D L F S F K I I I I W * R * G L T Y Q A
 Frame1 * S F F F * N Y Y Y L V E M R S H L S G
 DNA TGATCTTTTTCTTTAAAATTATTATTATTGGTAGAGATGAGGTCTCACTTATCAGGC 720
 | SNPG686 | SNPG667 hcv27463682 | SNPG710

 Frame3 L V S N S * S Q V I L P P Q P P K M Q G
 Frame2 G L K L L I S S D P P T S A S Q N A G I
 Frame1 W S Q T P D L K * S S H L S L P K C R D
 DNA TGGTCTCAAACCTCTGATCTCAAGTGATCCTCCACCTCAGCCTCCAAAATGCAGGGAT 780

 Frame3 L Q A * A T M P G P C C T G * D D C * V
 Frame2 T G M S H H A W S L L H W L G * L L G V
 Frame1 Y R H E P P C L V L V A L V R M T V R C
 DNA TACAGGCATGAGCCACCATGCCCTGGCTTGTGACTGGTAGGATGACTGTTAGGTGT 840
 | SNPG829

 Frame3 F K Q E * * E L T C L F T R N L N K F T
 Frame2 * T R M M R A H M F V Y K E L K Q I Y K
 Frame1 L N K N D E S S H V C L Q G T * T N L Q
 DNA TTAAACAAAGAATGATGAGAGCTCACATGTTTACAAGGAACCTAACAAATTACAA 900
 LF1.1 ATGATGAGAGCTCACATGTTTACAAGGAACCTAAA
 LF2.10 GCTCACATGTTTACAAGGAACCTAAA
 | SNPG856

 Frame3 R K K P I P I K K W A K D I N R H F S E
 Frame2 K K T H P H Q K V G K G Y K Q T L L R G
 Frame1 E K N P S P S K S G Q R I * T D T S Q R
 DNA GAAAAAAACCCATCCCCATCAAAAGTGGGCAAAGGGATATAAACAGACACTTCTCAGAGG 960
 | SNPG902 INS G DNA14AL1
 | SNPG908 INS A DNAP5AL1

 | SNP930 Pan troglod. A

3/12

FIGURE 1C

Frame3 E D I Y V A K K H M K K S S H T Y M K R
 Frame2 R H L R G Q E T Y E K K L T H V Y E T *
 Frame1 K T F T W P R N I * K K A H T R I * N V begin
 of homology orf1 to tre homologue <R154
 DNA AAGACATTTACGTGGCCAAGAAACATATGAAAAAGCTCACACACGTATATGAAACGTG 1020
 | End1stexon
 | SNP1006 | SNP1018

| SNPG972
 | R SNPG975

.....
 Frame3 D C L * S Y P K K N * F Q A T A V L L P
 Frame2 L F I I L S K K E L I S S N S S I T S I
 Frame1 T V Y N P I Q K R T D F K Q Q Q Y Y F H
 DNA ACTGTTTATAATCCTATCCAAAAAAGAACTGATTCAAGAACAGCAGTATTACTTCAT 1080
 | SNPG1048 | SNPG1056
 | SNP1069 | SNP1073
 | SNP1079

.....
 Frame3 F N T W T C K H Q K S H W R N * T T L *
 Frame2 Q Y L D L Q T S K K P L E K L N D S L K
 Frame1 S I L G P A N I K K A T G E T E R L S E
 DNA TCAATACTTGGACCTGCAAACATCAAAAAAGCCACTGGAGAAACTGAACGACTCTTGAA 1140

.....
 Frame3 K P * T K I * R S * N L E K T * G K G Q
 Frame2 A L N * D M K K L K S G K N L R K R T G
 Frame1 S L K L R Y E E V E I W K K L E E K D R
 DNA AGCCTTAAACTAAGATATGAAGAAGTTGAAATCTGAAAAACTTGAGGAAAAGGACAGG 1200
 R

hCV15819424 | SNPG1199
 | SNPG1181
 | SNPG1154

.....
 Frame3 A G G S T V A T T K K A G N R K R G W Q
 Frame2 R G K H S G Y N K K G R K Q E E R M A A
 Frame1 Q G E A Q W L Q Q K R Q E T G R E D G S
 DNA CAGGGGGAAAGCACAGTGGCTACAACAAAAAGGCAGGAAACAGGAAGAGAGGATGGCAGC 1260
 | SNPG1236

.....
 Frame3 H V G * R F F G D C I G F Q R Q N P K E
 Frame2 C W L K V L W R L Y W I P K T K P K R A
 Frame1 M L A K G S L E I V L D S K D K T Q K S
 DNA ATGTTGGCTAAAGGTTCTTGGAGATTGTATTGGATTCCAAGACAAAACCCAAAAGAGC 1320
 | SNPG1262 | R SNPG1263
 | SNPG1274 |
 | SNPG1319

.....
 Frame3 Q W * K E * K M * D Q R E R S N H S K G
 Frame2 M V K R M K N V R P K R K E Q S Q Q R N
 Frame1 N G E K N E K C E T K E K G A I T A K E
 DNA AATGGTAAAAAGAATGAAAATGTGAGACCAAAGAGAAAGGAGCAATCACAGCAAAGGAA 1380
 | R SNPG1334

.....

4/12

FIGURE 1D

Frame3 T I H N D D G * K H Q L D Y N G C S K N
 Frame2 Y T Q * * W I K T S A * L * W M L K E C
 Frame1 L Y T M M M D K N I S L I I M D A Q R M
 DNA CTATACACAATGATGATGGATAAAAACATCAGCTTGATTATAATGGATGCTCAAAGAATG 1440

| SNP14 SNPG1426

hcv27504712

.....
 Frame3 A G L S G F L Y F T F S Q C S * K S H Q
 Frame2 R I I R I P V F Y I L S V F L K K P S V
 Frame1 Q D Y Q D S C I L H S L S V P E K A I S
 DNA CAGGATTATCAGGATTCTGTATTTACATTCTCTAGTGTCCCTGAAAAAGCCATCAGT 1500
 K hcv11691030 |R SNPG1489

hcv16030281

| Y SNPG1466

| SNPG1444

| SNPG1487 SNP13

| SNPG1495 T

Pan trog.

.....
 Frame3 S R S H C * L D * S T P P R * F Y R Y M
 Frame2 Q E S L L A G L K H T S Q M I L * I H G
 Frame1 P G V T A S W I E A H L P D D D S I D T W
 DNA CCAGGAGTCACTGCTAGCTGGATTGAAGCACACCTCCAGATGATTCTATAGATACATGG 1560
 |SNPG1518
 | SNPG1505 hcv15819434
 | Y SNPG1508
 · | SNPG1521 |SNPG1554

.....
 Frame3 E E E G E C G V Y G T S * L V * F C K R
 Frame2 R R G G M W S I W Y F L T G L V L Q K I
 Frame1 K K R G N V E Y M V L L D W F S S A K D
 DNA AAGAAGAGGGGAATGTGGAGTATATGGTACTTCTGACTGGTTAGTTCTGCAAAAGAT 1620
 | SNPG1612 C Pan trog1.

.....
 Frame3 F T D W N N T L A S E R C T F Q V G K *
 Frame2 Y R L E Q H S G I * K M H F S S G K V K
 Frame1 L Q I G T T L W H L K D A L F K W E S K
 DNA TTACAGATTGGAACACACTCTGGCATCTGAAAGATGCACTTTCAAGTGGAAAGTAAA 1680
 |SNPG1630 hcv15819435

|SNPG1680

|SNPG1638 hcv15819436

.....
 Frame3 N C P V Q W A L G L W F * R E A I K T G
 Frame2 L S C A M G L G P L V L E G G Y K N W F
 Frame1 T V L C N G P W A F G F R G R L * K L V
 DNA ACTGTCTGTCAATGGCCTTGGCTTTGGTTAGAGGGAGGCTATAAAACTGGTT 1740
 |SNPGDELC17070
 |SNPG1687

5/12

FIGURE 1E

Frame3 S F A I P S I Q Q Q M L R S L H P H N T R
 Frame2 L C Y S Q Y T T N A K V T P P P P Q H Q N
 Frame1 P L L F P V Y N K C * G H S T P T T P E
 DNA CCTTGCTATTCCCAGTATAACAAATGCTAAGGTCACTCCACCCCCACAACACCAAGAA 1800
 | SNPG1742

.....
 Frame3 M K S C L S H W I L L I P H W K N Q F L
 Frame2 E E L S I S L D F T Y P S L E E S I P S
 Frame1 * R V V Y L I G F Y L S L I G R I N S F
 DNA TGAAGAGTTGCTATCTCATTGGATTTACTTATCCCTCATTGGAAGAATCAATTCTTC 1860
 |SNPG1810 Y
 |SNPG1813
 | SNPG1841

.....
 Frame3 L N L L P R C H L H L * K W M K T * N *
 Frame2 K P A A E M P P P P I K V D E D I E L I
 Frame1 * T C C R D A T S T Y K S G * R H R I D
 DNA TAAACCTGCTGCCGAGATGCCACCTCCACCTATAAAAGTGGATGAAGACATAGAATTGAT 1920
 | SNPG1890
 | SNPG1895 hcv16030290
 | SNPG1874
 | SNPG1875
 |SNP1907
 |SNP1909

.....
 Frame3 * V I K * V I M I K M R G Q D H * I Y Q
 Frame2 S D Q I S D N D Q N E R T G P L N I S I
 Frame1 K * S N K * * * S K * E D R T T E Y I N
 DNA AAGTGATCAAATAAGTGATAATGATCAAAATGAGAGGACAGGACACTGAATATATCAAT 1980
 | SNPG1957
 | SNPG1921
 | SNPG1922
 | SNPG1959
 | SNPG1976

.....
 Frame3 F Q L N Q L L L L N L M F H P S F S Q C
 Frame2 P V E S V A A S K S D V S P I I Q P V P
 Frame1 S S * I S C C F * I * C F T H H S A S A
 DNA TCCAGTTGAATCAGTTGCTGCTTAAATCTGATGTTCACCATCATTAGCCAGTGCC 2040
 | SNPG1992
 | SNPG1993

.....
 Frame3 L A * R M F H R L I I L K N W Q S N C L
 Frame2 S I K N V P Q I D H T K K L A V K L P E
 Frame1 * H K E C S T D * S Y * K T G S Q I A *
 DNA TAGCATAAAAGAATGTTCCACAGATTGATCATACTAAAAACTGGCAGTCAAATTGCCTGA 2100
 | SNPG2096

6/12

FIGURE 1F

Frame3 K S I * S N L K V Q I M S N S L L R M K
 Frame2 E H I I K S E S T N H E Q Q S P Q N E K
 Frame1 R A Y N Q I * K Y K S * A T V S S E * K
 DNA AGAGCATATAATCAAATCTGAAAGTACAAATCATGAGCAACAGTCTCCTCAGAATGAAAA 2160
 | SNPG2135

|| SNPG2110 del AA P. trogl. | EXON 3A IN P7N

.....
 Frame3 K L F L I V P P S Q * F P L Q L S C * Q
 Frame2 V I P D C S T K P V V S S P T L M L T D
 Frame1 S Y S * L F H Q A S S F L S N S H V N R
 DNA AGTTATTCCGTATTGTTCCACCAAGGCCAGTAGTTCTCTCCAATCTCATGTTAACAGA 2220
 | SNPG2180 hcv16030297
 | SNPG2192
 | SNPG2193 G Pan troglod.

.....
 Frame3 M K K R L I F M Q K L L F * W R K T N K
 Frame2 E E K A H I H A E T A L L M E K N K Q E
 Frame1 * R K G S Y S C R N C S S N G E K Q T R
 DNA TGAAGAAAAGGCTCATATTATGCAGAAACTGCTCTTAATGGAGAAAAACAAACAAGA 2280
 | SNPG2275

| SNPG2230

.....
 Frame3 K K N F R K D S K G N R K K L R R E E H
 Frame2 K E L Q E R Q Q G K Q K E frame shift of homology
 orf3 T E E G R T R
 Frame1 K R T S G K T A R E T E R N * G G K N T
 DNA AAAAGAACTTCAGGAAAGACAGCAAGGGAAACAGAAAGAAACTGAGGGAGGGAAAGAACACG 2340
 | SNPG2314

.....
 Frame3 E Q K A K K K Q E A E E N E I T Q K Q Q
 Frame2 A K S Q K E T R S * R K * N Y T E A T K
 Frame1 S K K P K R N K K L K M K L H R S N K
 DNA AGCAAAAAGCCAAAAAGAAACAAGAAGCTGAAGAAAATGAAATTACACAGAACAAAA 2400
 | SNPG2370
 | SNPG2375

.....
 Frame3 K A K E E M E K K E R E Q A K K E D K E
 Frame2 S K R R N G E E R T * T G Q E R G * R N
 Frame1 K Q K K K W R R K N V N R P R K R I K K
 DNA AAGCAAAAAGAAAGAAATGGAGAAGAAAGAACGTGAACAGGCCAAGAAAGAGGATAAAGAAA 2460
 | SNPG2449

hcv16030298

.....
 Frame3 I S A K K G K E I T R V K R Q S K S D H
 Frame2 L S K E G Q R N N K S K K T K * K * S *
 Frame1 S Q Q R R A K K * Q E * K D K V K V I M
 DNA TCTCAGCAAAGAAGGGCAAAGAAATAACAAGAGTAAAAGACAAAGTAAAGTGTATG 2520

.....T.....

7/12

FIGURE 1G

Frame3 E T S G A E K S V E D R G R R C S T P E
Frame2 N L W C R E V C R G Q G E K M F N P R S
Frame1 K P L V P R S L * R T G G E D V Q P Q K
DNA AACCTCTGGTGCCGAGAAGTCTGTAGAGGACAGGGGGAGAAAGATGTTCAACCCCAGAAG 2580
<EXE1626 <EXE1649
<EXE1631 not in Lucy <EXE1660

R2533

|SNPG2525

.....
Frame3 V Q K K S T R D V S H T S A T G D S G S
Frame2 T E K V N K R C V P Y I C D R G F R F R
Frame1 Y R K S Q Q E M C P I H L R Q G I Q V Q
DNA TACAGAAAAAGTCAACAAGAGATGTGTCCCATACATCTCGCACAGGGGATTCAAGGTTCAAG 2640
|SNPG2618 R
|SNPG2619 Y

FIGURE 2

9/12

FIGURE 3

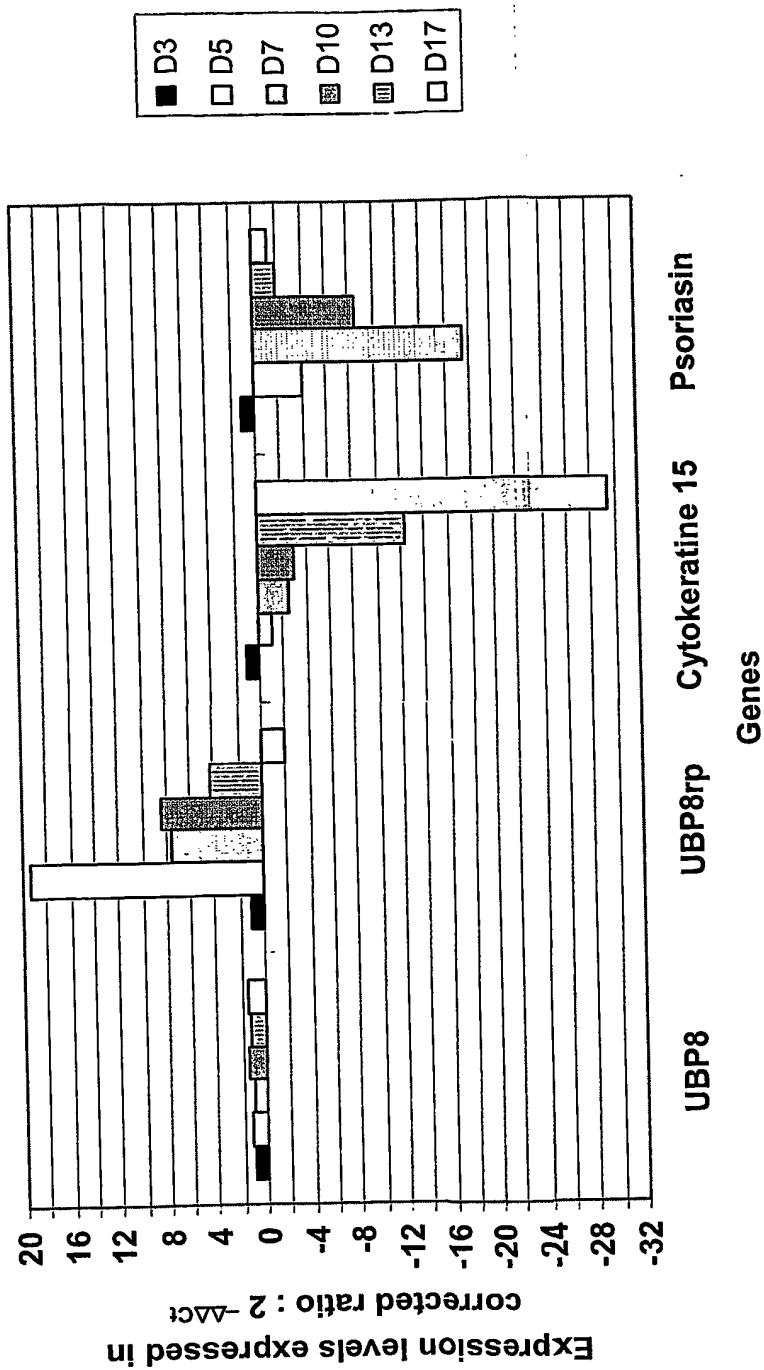
rhodanese *->tagelkalles.apkliliDvRspefGeryeyegGHIpAvNvpLee
SEQ ID NO:3 164 ta+el+ ++ +++ +li++D+ + + y+ ++I ++ vp e
TAKELYTMMMDkNISLIIMDAQRM-----QDYQDSCILHSLSVP-EK 204

rhodanese eiealldrsgilpdiekhlhllkdpeelaklfgelgsskdkrvivycrsgr
SEQ ID NO:3 205 +i+ +++s+i++ hl++d +++k+ g+ + + +s +
AISPGVTASWIEA-----HLPDDSIDTWKKRGNVEY---MVLLDWFS SAK 246

rhodanese dllrnrrsalaallkk1gypeVyiLkGGykeWlak--*
SEQ ID NO:3 247 + + ++ ++lk + kGGyk+W+ +
----DLQIGTTLWHLKDAL---FKWEKGGYKNWFLC 275

10/12

FIGURE 4



11/12

FIGURE 5

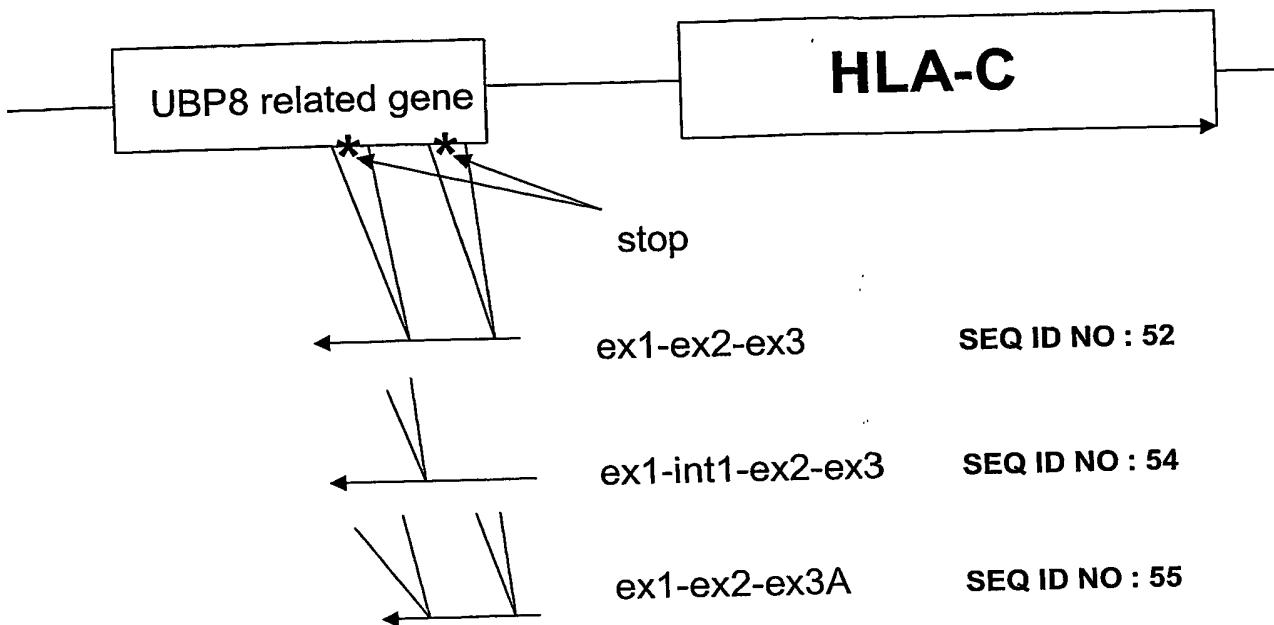


FIGURE 6

SEQ ID NO:3	MMRAHMFVYKELQIYKKKTHPHQKVKGKYQTLLRGRHRLRGQETYEKKLTHVYETPDFK
SEQ ID NO:53	MMRAHMFVYKELQIYKKKTHPHQKVKGKYQTLLRGRHRLRGQETYEKKLTHVYETPDFK
SEQ ID NO:56	MMRAHMFVYKELQIYKKKTHPHQKVKGKYQTLLRGRHRLRGQETYEKKLTHVYETPDFK *****
SEQ ID NO:3	QQQYYFHSILGPANIKKATGETERLSES LKLRYEEVEIWKKLEEKDRQGEAQWLQQKRQE
SEQ ID NO:53	QQQDCFRSILGPANIKKATGETERLSES LKLRYEEVEIWKKLEEKDRQGEAQWLLOOKRQE
SEQ ID NO:56	QQQDCFRSILGPANIKKATGETERLSES LKLRYEEVEIWKKLEEKDRQGEAQWLQQKRQE ***
SEQ ID NO:3	TGREDGSMILAKGSLEIVLDSKDQTOKSNGEKNEKCETKEKGAIATAKELYTMMMDKNISLI
SEQ ID NO:53	TGREDGSTLAQDSLEIVLDSKDQTOKSNGEKNEKCETKEKGAIATAKELYTMMMDKNISLI
SEQ ID NO:56	TGREDGSTLAQDSLEIVLDSKDQTOKSNGEKNEKCETKEKGAIATAKELYTMMMDKNISLI *****
SEQ ID NO:3	IMDAORMQDYQDSCILHSLSVPEKAISP GVTASWIEAHLPPDDSDITWKKRGNVEYMVL LD
SEQ ID NO:53	IMHAORMQYYQDSCILHSLSVPEKAISP GVTASWIEAHLPPDDSDITWKKRGNVEYMVL LD
SEQ ID NO:56	IMHAORMQYYQDSCILHSLSVPEKAISP GVTASWIEAHLPPDDSDITWKKRGNVEYMVL LD ***
SEQ ID NO:3	WFSSAKDLQIGTTLWHLKDALKWEKGGYKNWFLCYSQYTTNAKVTPPPQHQNEELSISL
SEQ ID NO:53	WFSSAKDLQIGTTLWHLKDALKWEKGGYKNWFFCYSQYTTNAKVTPPPQHQNEELSISL
SEQ ID NO:56	WFSSAKDLQIGTTLWHLKDALKWEK----- *****
SEQ ID NO:3	DFTYPSLEESIPS KPAAE MPPPPIKVDEDIELISDQISDNDQNERTGPLNISIPVESVAA
SEQ ID NO:53	DFTYPSLEESIPS KPAAE MPPPPIEVDEDIELISDQISDNDQNERTGPLNISIPVESVAA
SEQ ID NO:56	----- *****
SEQ ID NO:3	SKSDVSP IIQP VPSIKNVPQIDHTKKLAVKLPEEHIIKSESTNHEQQSPQNEKVIPDCS A
SEQ ID NO:53	SKSDVSP IIQP VPSIKNVPQIDHTKKLAVKLPEEHIIKSESTNHEQQSPQNEKVIPDCS A
SEQ ID NO:56	-----SPQNEKVIPDCS A *****:
SEQ ID NO:3	KPVSSPTLMLTDEEKAHIAETALLMEKNQKEKELQERQQGKQKETEEGRTRAKSQET
SEQ ID NO:53	KPVSSPTLMLTDEEKAHIAETALLMEKNQKEKELQERQQGKQKETEEGRTRAKSQET
SEQ ID NO:56	KPVSSPTLMLTDEEKAHIAETALLMEKNQKEKELQERQQGKQKETEEGRTRAKSQET *****
SEQ ID NO:3	RS---
SEQ ID NO:53	RSCRK
SEQ ID NO:56	RSCRK **